



(11) **EP 1 845 156 A1**

(12) **EUROPEAN PATENT APPLICATION**
published in accordance with Art. 158(3) EPC

(43) Date of publication:
17.10.2007 Bulletin 2007/42

(21) Application number: **06702529.6**

(22) Date of filing: **05.01.2006**

(51) Int Cl.:
C12N 15/00 (2006.01) **C12N 15/09** (2006.01)
C12Q 1/02 (2006.01) **C12Q 1/68** (2006.01)
G01N 33/53 (2006.01)

(86) International application number:
PCT/JP2006/300242

(87) International publication number:
WO 2006/080192 (03.08.2006 Gazette 2006/31)

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IS IT LI LT LU LV MC NL PL PT RO SE SI
SK TR**

(30) Priority: **05.01.2005 JP 2005001033**

(71) Applicant: **National University Corporation Chiba
University
Chiba-shi
Chiba, 2638522 (JP)**

(72) Inventors:
• **SHIMADA, Hideaki,**
c/o Graduate School of Medicine
Chiba-shi, Chiba 2608670 (JP)
• **TOMONAGA, Takeshi,**
c/o Graduate School of Medicine
Chiba-shi, Chiba, 2608670 (JP)

- **MATSUSHITA, Kazuyuki,**
Graduate School of Medicine
Chiba-shi,
Chiba 2608670 (JP)
- **OCHIAI, Takenori,**
c/o Graduate School of Medicine
Chiba-shi, Chiba, 2608670 (JP)
- **NOMURA, Fumio,**
c/o Graduate School of Medicine
Chiba-shi, Chiba, 2608670 (JP)

(74) Representative: **Dunleavy, Kevin James**
Knoble & Yoshida,
p/o De Vries & Metman,
Overschiestraat 180
1062 XK Amsterdam (NL)

(54) **GENE SPECIFIC TO CANCER AND DIAGNOSIS KIT USING THE SAME**

(57) The present invention provides a new additional identification of a gene related to cancer expression and a diagnostic kit using the same

EP 1 845 156 A1

Description

FIELD OF THE INVENTION

5 [0001] The present invention relates to a cancer-specific gene and a diagnostic kit using the same, and further relates to a method for using the diagnostic kit.

BACKGROUND OF THE INVENTION

10 [0002] The most crucial challenge in measures against cancer is the early detection of cancer. Particularly, early detection is important for cancers originating from the upper part of the large intestine since they cause only limited subjective symptoms and the medical condition may likely to be in its advanced stage by the time of discovery.

[0003] Traditional measures against large intestine cancer include screening by the fecal occult blood test, diagnosis by serum markers such as CEA or CA19-9, and diagnosis during a course of treatment. However, the positive rates of these methods are high only for advanced cancers and extremely low for early cancers, making accurate diagnosis difficult in their early stages.

[0004] Meanwhile, a biological diagnostic method using a cancer tissue-specific protein marker is suggested as a method allowing simple and reliable early diagnosis of malignancy. This method can be performed on a broad range of asymptomatic subjects since it does not require a large-scale facility and causes small burdens for the subject. For instance, Japanese Patent Application Publication No. H07-51065 discloses a usage of glycoprotein 39 as a tumor marker.

[0005] In addition, International Patent Application Publication No. WO/2004/018679 describes a technology regarding a cancer diagnostic kit using CENP-A.

[0006] However, cancer expression cannot be thoroughly verified by the technology described in the above JP-A-H7-51065 alone and a plurality of means must be used to ensure positive identifications.

25 [0007] Considering the above situation, the purpose of the present invention is to provide a further identification of a gene related to cancer expression and a diagnostic kit using the same.

SUMMARY OF THE INVENTION

30 [0008] Considering the above situation, the present invention employs specific means described below.

[0009] A first means is a polynucleotide as in one of the following (a)-(c):

A polynucleotide consisting of a base sequence as in SEQ ID NO: 1 or a complementary base sequence thereof;
 (b) a polynucleotide consisting of a base sequence having at least 70% homology with the base sequence as in
 35 SEQ ID NO: 1 or the complementary base sequence thereof; and (c) a polynucleotide coding a protein consisting of an amino-acid sequence as in SEQ ID NO: 2, or another amino-acid sequence defined by the amino-acid sequence having one or several amino acid deletions, substitutions or additions.

[0010] This means is preferably used as a marker for detecting a cancer. The polynucleotide in this means may greatly contribute to a cancer diagnosis when used as a marker since the polynucleotide has been discovered to exhibit a high expression in cancer tissues. The homology with the base sequence as in SEQ ID NO: 1 is preferably equal to or greater than 70%, more preferably equal to or greater than 80%, and even more preferably equal to or greater than 90%.

[0011] A second means is a nucleic acid molecule including the polynucleotide as in the first means.

[0012] A third means is a cancer diagnostic kit using a factor specific to the polynucleotide as in the first means. The term "specific gene" as used herein refers to a gene having a significantly higher affinity to a polynucleotide in question than that to other irrelevant polynucleotides (particularly the ones with less than 30% identity). The affinity can be measured by, for example, hybridization assay or binding assay.

[0013] In the above means, the specific factor is preferably at least one of the group consisting of nucleic acid molecules, polypeptides, lipids, carbohydrate chains, low-molecular-weight organic molecules, and complex molecules thereof; and
 50 a cancer diagnosed by the cancer diagnostic kit is preferably a rectal cancer or a colon cancer.

A forth means is a primer consisting of a base sequence as in SEQ ID NO: 3.

A fifth means is a primer consisting of a base sequence as in SEQ ID NO: 4.

55 A sixth means is a primer set consisting of the primers as in SEQ ID Nos: 3 and 4.

[0014] A seventh means is a method comprising the steps of: measuring an expression level of a protein consisting

of an amino-acid sequence as in SEQ ID NO: 2 for each of two collected cells; and calculating a ratio between the measured expression levels. In this case, one of the two samples is collected from a non-cancer tissue; the other sample is collected from, for example, a tissue suspected to be a cancer tissue; and if the expression level ratios of the two samples are different, the subject from whom the samples were collected may be determined to be at a high risk of cancer.

[0015] Preferably for this means, the step of measuring the expression level of the protein is performed with the western blot, wherein one of the two collected cells is a cell in a non-cancer tissue, wherein one of the two collected cells is a cell in a cancer tissue; and the step of measuring the expression level further comprises the step of determining whether or not the ratio between the measured expression levels is equal to or greater than 1.7.

[0016] As described above, a gene related to cancer expression may be newly identified, and a diagnostic kit using the same may be provided.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017]

Fig. 1 is a diagram showing western blot results for CENP-H;

Fig. 2 is a diagram showing western blot results for hMis12;

Fig. 3 is a diagram showing the results of staining cross-sectional areas of rectal cancer tissues and their respective nearby non-rectal cancer tissues with anti-human CENP-H polyclonal antibody;

Fig. 4 is a diagram showing analysis results on amounts of CENP-H mRNA in each of the surfaces of rectal cancer tissues and normal tissues using RT-PCR and real-time quantitative PCR; and

Fig. 5 is a diagram showing analysis results on amounts of CENP-H mRNA in each of the surfaces of rectal cancer tissues and normal tissues using RT-PCR and real-time quantitative PCR.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0018] One embodiment of the present invention will be described below.

Tissue Collection

[0019] Body tissues were collected with a surgical method from 15 patients with early-stage colorectal cancer. The tissues were taken from cancer tissue (hereinafter referred to as "cancer tissue") and tissue at a part 5-10 cm away from the cancer tissue (hereinafter referred to as "non-cancer tissue"), respectively. The collected tissues were immersed in liquid nitrogen and stored at -80°C.

Protein Extraction

[0020] The cryonically-preserved tissues were then placed into lysis buffer (7M urea, 2M thiourea, 2% 3-[(3-Cholamidopropyl)Dimethylammonio]-1-Propanesulfonate, 0.1M DTT, 2% IPG buffer (made by Amersham Pharmacia Biotech), 40mM Tris), lysed using polytron homogenizer (made by Kinematica), and centrifuged at 10,000g and 4°C for 1 hour to collect supernatant and extract proteins.

Immunoblot

[0021] The proteins were blotted to polyvinylidene fluoride membranes (made by Millipore) in a tank transfer device (made by Bio-Rad) and the membranes were then blocked with phosphate buffered saline (PBS) containing 5% skim milk. Next, 1:5000 diluted rabbit anti-CENP-H antibody, 1:100 diluted rabbit anti-hMis12 antibody and 1:500 diluted goat anti-β-actin antibody, each placed in the blocking buffer, were used as a primary antibody; and 1:3000 diluted goat anti-rabbit IgG HRP and 1:500 diluted rabbit anti-goat IgG HRP, each placed in the blocking buffer, were used as a secondary antibody.

[0022] Note that antibodies on the antigen membrane were detected with enhanced chemiluminescence detection reagent (made by Amersham Pharmacia Biotech). Also the intensity of each band was measured with an NIH image.

PCR and Real-Time Quantitative PCR

[0023] TotalRNA was extracted from the cancer tissue and the non-cancer tissue, respectively, using RNeasy Mini Kit (made by Qiagen). Also cDNA was synthesized from each extracted totalRNA, respectively, using a 1st Strand cDNA Synthesis Kit for RT-PCR (made by Roche).

[0024] Then each cDNA obtained by this synthesis was used as a template to amplify the cDNA of CENP-H with PCR. In the PCR, a primer comprising a base sequence as in SEQ ID NO: 3 or 4 were used as the forward and reverse primer, respectively, and cDNA of GAPDH or β -actin were amplified as the controls.

[0025] Subsequently cDNA real-time quantitative PCR for CENP-H was performed in a LIGHTCYCLER®capillary. For the PCR reaction mixture, 3.0mM of $MgCl_2$, 0.5 μ M of the primer as in SEQ ID NO: 3 and 0.5 μ M of the primer as in SEQ ID NO: 4 were added to LIGHTCYCLER® DNA Master SYBR Green I (FastStart Taq DNA polymerase, dNTP, buffer, SYBR Green I), and the procedure was conducted within a total of 2.0 μ l.

[0026] LIGHTCYCLER® software version 3.3 (made by Roche) was then used for analysis.

Immunohistochemical Staining Method

[0027] The frozen tissue sections were dried on a glass slide and fixed in 4°C acetone. The tissues were then washed with PBS 3 times and blocked with the blocking buffer (10% fetal bovine serum/PBS) for 1 hour.

[0028] The sample was incubated for 1 hour in 3% bovine serum albumin/PBS using one or both of 1:2000 diluted rabbit anti-CENP-H antibody and 1:1000 diluted anti-human CENP-A monoclonal antibody. After washing with PBS, the sample was incubated for 1 hour with 1:1000 diluted ALEXA FLUOR® 488- or 594-bound goat anti-rabbit anti-mouse IgG secondary antibody (made by Molecular Probes) and/or ALEXA FLUOR® 594-bound goat anti-mouse IgG secondary antibody.

[0029] DNAs were counterstained using DAPI III Counterstain (made by Vysis). The sample was observed with a fluorescence microscope (made by Leica QFISH). The tissue sections were stained with hematoxylin for 30 minutes for HE staining, dried over 100% ethanol and xylene and encapsulated with Permount.

Results

[0030] Fig. 1 shows results of the western blot. As shown in Fig. 1, CENP-H was highly expressed in the cancer tissue in any of the 15 cases. Particularly, a ratio between the non-cancer tissue and the cancer tissue CENP-H expressions was 1.7-9.6, indicating a large difference between these two kinds of tissues. For another centromere protein hMis12, on the other hand, no notable difference was discovered between the cancer tissue and the non-cancer tissue. (See Fig. 2, in which tissues with the same case number as in Fig. 1 are identical with those in Fig. 1).

[0031] Next, in order to verify that CENP-H is expressed in the cancer cell, but not in stromal cells, cross-sectional areas of colorectal cancer tissues and nearby non-cancer tissues were stained with an anti-human CENP-H polyclonal antibody. The results are shown in Fig. 3. Note that Fig. 3 (a) shows an HE-stained image of the cancer tissue; Fig. 3 (b), (c) and (d) show a CENP-H antibody immunostained image of the cancer tissue; Fig. 3 (e) shows an HE-stained image of the non-cancer tissue; and Fig. 3 (f) shows a CENP-H-stained image of the non-cancer tissue.

[0032] As a result, it was confirmed that CENP-H existed as small patchy points in cell nuclei at positions coinciding with the centromeres in a similar manner to that of other centromere proteins such as CENP-A and CENP-C. It was also confirmed that the CENP-H had been increased both in number and size in the cancer tissues (Fig. 3 (c) and (d)) compared to the non-cancer tissue (Fig. 3 (f)). It should be noted that the stained CENP-H was verified not in the stromal cells, but in the cancer tissue epithelia. Also the present experiment was conducted on various tissue sections and all the results were similar to each other.

[0033] Accordingly, it was confirmed that CENP-H was expressed in cancer cells.

[0034] Subsequently, in order to verify that the CENP-H overexpression was a result of its increase due to transcription, amounts of mRNA of CENRH in the colorectal cancer tissues and the non-cancer tissues were analyzed, respectively, using RT-PCR and real-time quantitative PCR. The results are shown in Figs. 4 and 5.

[0035] As shown in Fig. 4, an expression level of the CENP-H mRNA in the cancer tissues is far more increased than in the non-cancer tissues. Furthermore, it was discovered that the expression level of the CENP-H mRNA indicated a strong correlation with the CENP-H expression ratio between the non-cancer tissues and the cancer tissues illustrated in Fig. 1. As a control, when confirmation was done pertaining to GAPDH, there was no significant difference in the expression level of the CENP-H mRNA between the cancer tissues and the non-cancer tissues.

[0036] Fig. 5 illustrates a comparison of the mRNA expression levels shown in Fig. 4 between the non-cancer tissues and the cancer tissues using StatView statistical analysis software. According to Fig. 5, the CENP-H mRNA expression level in the cancer tissues (Cancer) was 5 times higher than that of the non-cancer tissues (Normal). In this way, the presence of cancer can be also verified by examining the CENP-H expression level.

Industrial Availability

[0037] Thus according to the present invention, a gene related to cancer expression may be newly identified, and a diagnostic kit using the same may be also provided.

5

10

15

20

25

30

35

40

45

50

55

<160>4

<211>744

<212>DNA

atggaggagc agccccagat gcaagacgcc gacgagcccg cggaciccgg agggggaaggc	60
cgggcaggcg ggccaccgca ggtcgccggc gcccaggcgg cgtgcagcga ggaccgcatg	120
acctgctcc tcaggctgag agcacagaca aaacaacaac tcttagaata taaatcaatg	180
gttgatgcaa gtgaagaaaa aactccagaa caaattatgc aagaaaaagca aatcgaagct	240
aaaattgaag acctggaaaa tgaaattgaa gaggtaaaag ttgcttttga gataaaaaag	300
cttgcattag acaggatgag acttcaact gcacttaaaa aaaacctgga gaaaaattagc	360
agacagtcta gtgtgctcat ggataacatg aaacacctat tagagctaaa taaattaata	420
atgaaatcac agcaggaatc ttgggattta gaggaaaaac tgcttgatat tagaaaagaag	480
agattgcaat taaaacaagc ttcagaaaat aagcttttag aaatacagac tgaagaagaac	540
aaacagaaga ttgatttgga cagtatggaa aactcagaga ggataaagat catacgacaa	600
aacctacaga tggagataaa aattactact gtatttcaac atgtgttcca gaaccttatt	660
ttggggagta aagtcaattg ggagaggat cctgccctta aggaattgt tctgcagctt	720
gagaagaatg ttgacatgat gtaa	744

 $\langle 210 \rangle_2$

<211>247

<212>PRT

<213>Homo sapiens

<400>

Met Glu Glu Gln Pro Gln Met Gln Asp Ala Asp Glu Pro Ala Asp Ser

1 5 10 15

Gly Gly Glu Gly Arg Ala Gly Gly Pro Pro Gln Val Ala Gly Ala Gln

20 25 30

Ala Ala Cys Ser Glu Asp Arg Met Thr Leu Leu Leu Arg Leu Arg Ala

35 40 45

Gln Thr Lys Gln Gln Leu Leu Glu Tyr Lys Ser Met Val Asp Ala Ser

50

EP 1 845 156 A1

Glu Glu Lys Thr Pro Glu Gln Ile Met Gln Glu Lys Gln Ile Glu Ala
 65 70 75 80
 5 Lys Ile Glu Asp Leu Glu Asn Glu Ile Glu Glu Val Lys Val Ala Phe
 85 90 95
 Glu Ile Lys Lys Leu Ala Leu Asp Arg Met Arg Leu Ser Thr Ala Leu
 100 105 110
 10 Lys Lys Asn Leu Glu Lys Ile Ser Arg Gln Ser Ser Val Leu Met Asp
 115 120 125
 Asn Met Lys His Leu Leu Glu Leu Asn Lys Leu Ile Met Lys Ser Gln
 15 130 135 140
 Gln Glu Ser Trp Asp Leu Glu Glu Lys Leu Leu Asp Ile Arg Lys Lys
 145 150 155 160
 20 Arg Leu Gln Leu Lys Gln Ala Ser Glu Ser Lys Leu Leu Glu Ile Gln
 165 170 175
 Thr Glu Lys Asn Lys Gln Lys Ile Asp Leu Asp Ser Met Glu Asn Ser
 180 185 190
 25 Glu Arg Ile Lys Ile Ile Arg Gln Asn Leu Gln Met Glu Ile Lys Ile
 195 200 205
 Thr Thr Val Ile Gln His Val Phe Gln Asn Leu Ile Leu Gly Ser Lys
 30 210 215 220
 Val Asn Trp Ala Glu Asp Pro Ala Leu Lys Glu Ile Val Leu Gln Leu
 225 230 235 240
 35 Glu Lys Asn Val Asp Met Met
 245
 40 <210>1
 <211>20
 <212>DNA
 <213>Homo sapiens
 45 <400>cagtcctagtg tgctcatgga t 21
 <210>1
 50 <211>20
 <212>DNA
 <213>Homo sapiens
 55 <400>tccatctgta ggtttgtcg 20

Claims

1. A polynucleotide as in one selected from the group consisting of the following (a)-(c):

- (a) a polynucleotide consisting of a base sequence as in SEQ ID NO: 1 or a complementary base sequence thereof;
- (b) a polynucleotide consisting of a base sequence having at least 70% homology with the base sequence as in SEQ ID NO: 1 or the complementary base sequence thereof; and
- (c) a polynucleotide coding a protein consisting of an amino-acid sequence as in SEQ ID NO: 2, or another amino-acid sequence defined by the amino-acid sequence having one or several amino acid deletions, substitutions or additions.

2. The polynucleotide as in Claim 1, wherein the polynucleotide is a marker for detecting a cancer.

3. A nucleic acid molecule including the polynucleotide as in Claim 1.

4. A cancer diagnostic kit using a factor specific to the polynucleotide as in Claim 1.

5. The cancer diagnostic kit as in Claim 4, wherein the specific factor is at least one of the group consisting of nucleic acid molecules, polypeptides, lipids, carbohydrate chains, low-molecular-weight organic molecules, and complex molecules thereof.

6. The cancer diagnostic kit as in Claim 5, wherein a cancer diagnosed by the cancer diagnostic kit is a rectal cancer or a colon cancer.

7. A primer consisting of a base sequence as in SEQ ID NO: 3.

8. A primer consisting of a base sequence as in SEQ ID NO: 4.

9. A primer set consisting of the primer as in Claim 7 and the primer as in Claim 8

10. A method including the steps of: measuring an expression level of a protein consisting of an amino-acid sequence as in SEQ ID NO: 2 for each of two collected cells; and calculating a ratio between the measured expression levels.

11. The method as in Claim 10, wherein the step of measuring the expression level of the protein is performed with the western blot.

12. The method as in Claim 10, wherein one of the two collected cells is a cell in a non-cancer tissue.

13. The method as in Claim 10, wherein one of the two collected cells is a cell in a cancer tissue.

14. The method as in Claim 10, further including the step of determining whether or not the ratio between the measured expression levels is equal to or greater than 1.7.

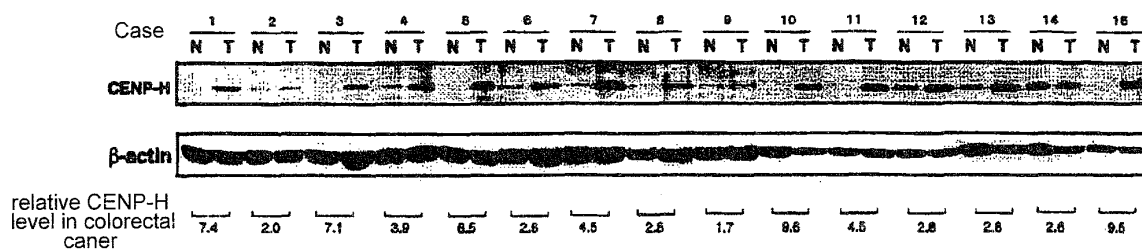


FIG. 1

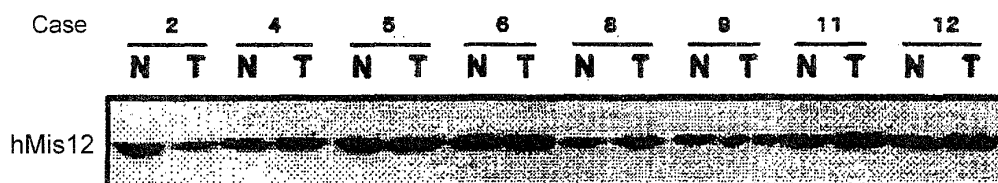


FIG. 2

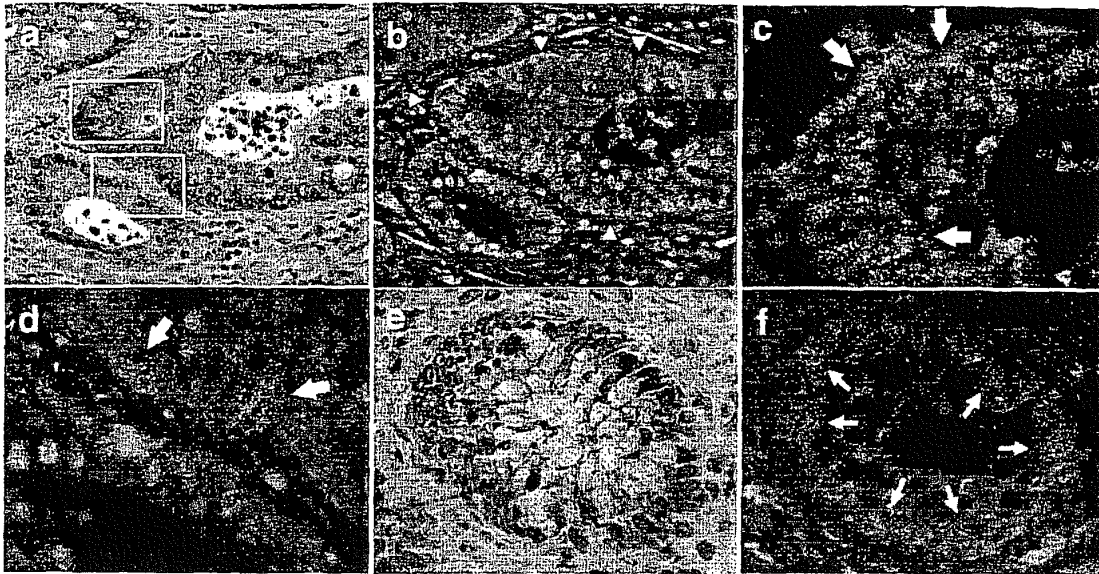


FIG. 3

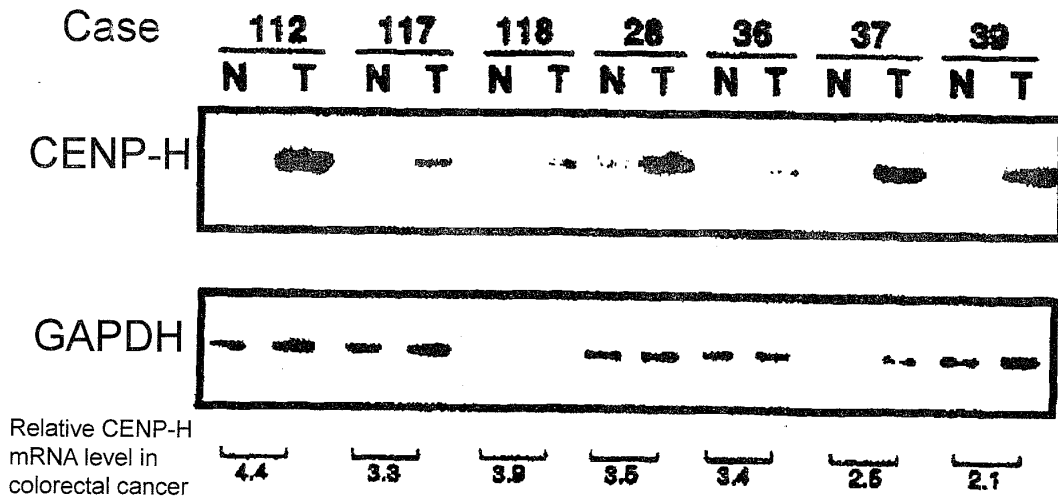


FIG. 4

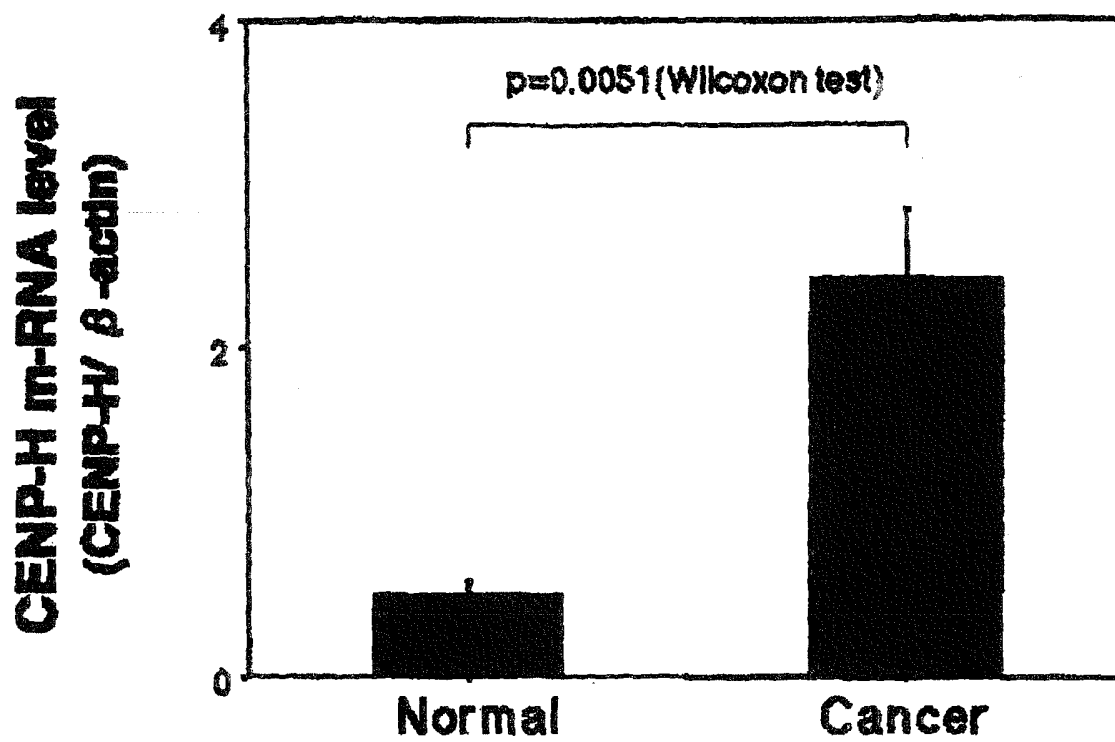


FIG. 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2006/300242

A. CLASSIFICATION OF SUBJECT MATTER C12N15/00 (2006.01), C12N15/09 (2006.01), C12Q1/02 (2006.01), C12Q1/68 (2006.01), G01N33/53 (2006.01) According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12N15/00 (2006.01), C12N15/09 (2006.01), C12Q1/02 (2006.01), C12Q1/68 (2006.01), G01N33/53 (2006.01) Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2006 Kokai Jitsuyo Shinan Koho 1971-2006 Toroku Jitsuyo Shinan Koho 1994-2006 Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) BIOSIS/MEDLINE/WPIDS (STN), GenBank/EMBL/DDBJ/GeneSeq, SwissProt/PIR/Geneseq, JSTPlus(JDream2)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Sugata N. et al., "Human CENP-H multimers colocalize with CENP-A and CENP-C at active centromere-kinetochore complexes", (2000), Hum.Mol.Genet., Vol.9, No.19, pages 2919 to 2926	1, 3, 7-9
X	Sugata N. et al., "Characterization of a novel kinetochore protein, CENP-H", (1999), J.Biol.Chem., Vol.247, No.39, pages 27343 to 27346	1
Y	WO 2004/018679 A1 (Japan Science and Technology Agency), 04 March, 2004 (04.03.04), (Family: none)	2, 4-6, 10-14
☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 12 April, 2006 (12.04.06)		Date of mailing of the international search report 16 May, 2006 (16.05.06)
Name and mailing address of the ISA/ Japanese Patent Office		Authorized officer
Facsimile No.		Telephone No.

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2006/300242

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2003/104426 A1 (Merck & Co., Inc.), 18 December, 2003 (18.12.03), & EP 1529055 A1 & US 2005/0164201 A1	2, 4-6, 10-14
P, X	Tomonaga T. et al., "Centromere protein H is up-regulated in primary human colorectal cancer and its overexpression induces aneuploidy", (2005, Jun.), Cancer Res., Vol.65, No.11, pages 4683 to 4689	1-14

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- JP H0751065 B [0004]
- WO 2004018679 A [0005]
- JP H751065 A [0006]